

How language helps us think

RAY JACKENDOFF

In memory of John Macnamara

On formal and empirical grounds, the overt form of language cannot be the vehicle that the mind uses for reasoning. Nevertheless, we most frequently experience our thought as "inner speech". It is argued that inner speech aids thought by providing a "handle" for attention, making it possible to pay attention to relational and abstract aspects of thought, and thereby to process them with greater richness. Organisms lacking language have no modality of experience that provides comparable articulation of thought; hence certain kinds of thought very important for human intelligence are simply unavailable to them.

1. Introduction

The influence of language on thought seems so simple and natural as to require no explanation. In this paper, therefore, I must first show that the phenomenon *does* require an explanation, before being able to give an idea of what the explanation might be like.

Here's the problem. It is a widespread platitude that we differ from other animals in being smarter, in being able to think (or reason) better. It is another widespread platitude that we differ from other animals in having language. Hence it is not surprising to draw some connection between these two: the connection between language and thought seems altogether obvious, not even worth inquiring about. *Of course* language helps us think.

But how? In answering, I wish to avoid a narrow human chauvinism, by assuming evolutionary continuity between us and the animals. Thus the question may be focused further: How much of our increased reasoning power

is due to our big brains, and how much of it is specifically due to the presence of a language faculty?

From the way I am putting these questions, the reader may note that I am emphatically not expecting an absolute answer; otherwise I would have entitled the paper something like "Does language enable us to think?" Rather, I will be promoting a somewhat more complex position:

- (1) Thought is a mental function completely separate from language and can go on in the absence of language.
- (2) On the other hand, language provides a scaffolding that makes possible certain varieties of thought that are more complex than are available to nonlinguistic organisms.

As a good cognitive scientist, I want ultimately to explain every step of this position, with no appeals to intuition — or magic — at any point. So an underlying part of the question is: how in the world does the brain accomplish thought, and how does it produce and understand language?

In approaching these questions, it's important to avoid falling into a common trap (of which I probably don't need to remind you). The classic illustration comes from the problem of vision: how do we see things and know what we are seeing? The science show on television tells us that an image of the world is projected onto the retina, which transmits this image to the brain, where we see it. The problem is, who's looking at the image in the brain? There's no little person (or homunculus) in there looking at the image to understand it. How would the homunculus see? That would just push the problem back one level (and threaten an infinite regress). No, in the brain there are just neurons, each of which just responds to firings impinging on it by itself firing, thereby stimulating other neurons. Somehow all those firings have to add up to our experience of seeing. This point has been forcefully pushed home in recent years by writers as diverse as Dennett (1991), Crick (1994), and Edelman (1992).

Similarly with language: there's no one in there listening, there are just more neurons firing. And with reasoning: we don't understand things because there's a little person in there doing our understanding. What we experience as understanding must again be just a product of neurons firing.

In order to keep ourselves honest, not to appeal to magic, it's often useful to think of the brain as a sort of biological computer whose program we have to discover. Computers have no common sense, so every step underlying behavior and experience has to be absolutely explicit. This is not to say, of course, that the brain computes along principles at all similar to those of your PC — in fact the traditional Turing machine analogy is fortunately now open to serious question, thanks to the last decade's flourishing exploration of neural net-style computation. Can we do better than a computational approach, and explain language and reasoning really in terms of the neurons? Searle (1992) and Edelman (1992), among others, have offered blistering attacks on computational approaches to brain function, arguing that only in an account of the neurons is an acceptable theory to be found. I agree that eventually the neuronal basis of mental functioning is a necessary part of a complete theory. But current neuroscience, exciting as it is, at the moment is far from being able to tackle the question of how language helps us think. For instance:

1. Through animal research, brain imaging techniques, and studies of brain damage, we know what many areas of the brain do as a whole. But with the exception of certain lower-level visual areas, we have little idea *how* those areas do what they do. (It is like knowing what parts of a television set carry out what functions, without knowing really how those parts work.)
2. We know how certain neurotransmitters affect brain function in a global way, which may explain certain overall effects or biases such as Parkinson's disease or depression. But this doesn't inform us about the fine-scale articulation of brain function, say how the brain stores and retrieves individual words.
3. We know a lot about how individual neurons and some small systems of neurons work, but again little is known about how neurons encode words, or even speech sounds — though we know they must do it somehow.

So, for the moment the computer analogy is among the best tools we have for understanding the brain at the level of functioning relevant to language and thought, and over the years it has proven a pragmatically useful perspective.¹

2. Brain phenomena opaque to awareness

In working out a computational theory of brain function, it immediately becomes necessary to claim that certain steps behind common sense are hidden from conscious experience. Such a claim supposes that we can distinguish two sorts of phenomena in the brain: those that are present to awareness — that show their face as something we experience — and those that are not (with perhaps some in-between cases as well). What can it mean to say a brain phenomenon is not present to awareness, that it is unconscious? Freud, of course, shocked the world by claiming that our behavior is driven in part by unconscious and uncomfortably dark motives. But he thought that these motives could be brought to light by suitable introspection, by psychotherapy. Cognitive science actually makes a stronger claim: some phenomena in the brain are *never* present to awareness, no matter how much we try.

For a simple example, notice how the images in your two eyes are subtly different. You can alternately close one eye then the other, and see the image change. But in normal binocular vision, the two images are fused into a single image, which contains an element of depth not present in either eye separately. How does your brain do that? No matter how hard you think about it, you can't catch yourself in the act of fusing the two images into one. Phenomenologically, it just happens, as if by magic. Explaining exactly how the brain does this is a major preoccupation of people in vision research (e.g., Julesz 1971, Marr 1982, Crick 1994). But even when they figure out how the brain does it, they still won't be able to catch themselves in the act! And this is true of the peripheral processes of perception in general. We can't catch our auditory system in the process of performing a frequency analysis of sounds; we just hear a sound (Bregman 1990). We can't catch our proprioceptive system in the process of figuring out how heavy a load we're carrying by detecting the stretch of the muscles; we just feel the weight (Lackner 1985).

In fact, it seems plausible that all peripheral sensory processes in the brain are totally inaccessible to awareness. A crude diagram like Figure 1 schematizes the situation: what we might call the "outer ring" of brain processes is completely unconscious.

Once we've established that *some* brain phenomena are unconscious, the obvious question arises as to exactly *which* are unconscious. Is it just the periphery, or is there more? Section 4 will try to show that there is indeed great deal more, and subsequent sections will work out interesting consequences for the relation of language and thought.

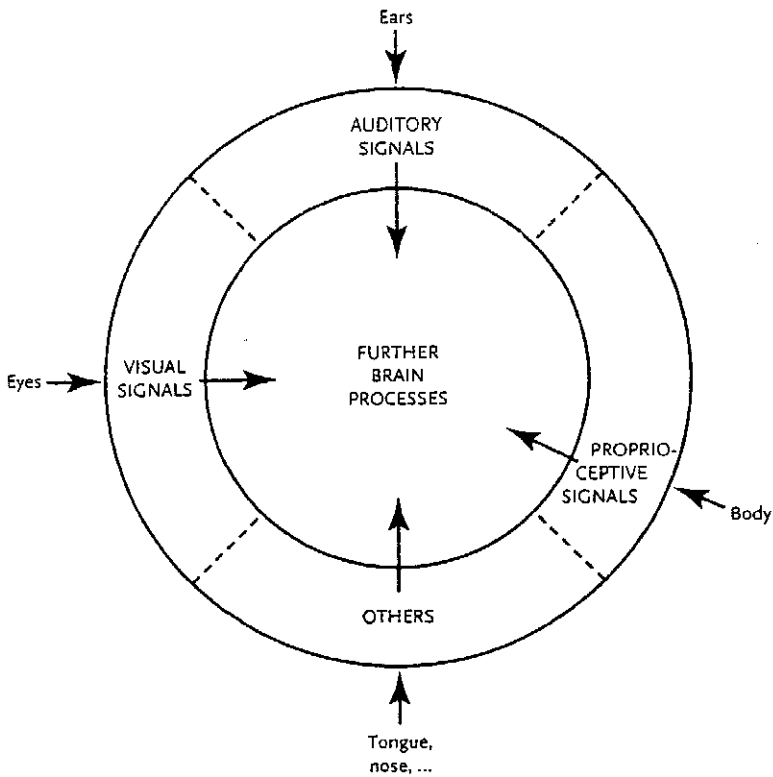


Figure 1

Meanwhile, the reader may notice that I am not asking the questions everyone seems to ask, namely “What is consciousness?” or “What is consciousness for?” (This includes, for instance, Hofstadter 1979, Searle 1992, Baars 1988, and Block 1995.) I suspect that many people who insist on this question (though perhaps not the authors just cited) really want to know something like “What makes human life meaningful?”. And in many cases — my students have often been quite explicit about this — they are just hoping that somehow we’ll be able to bring back some magic, some hope for a soul or the possibility of immortality. By contrast, I am asking a less cosmic and more structural question, simply, “Which brain phenomena appear in consciousness and which ones do not, and what are the consequences for the nature of experience?”. If you don’t care for this question, I am the first to acknowledge

that it's a matter of taste; but I hope to convince you that my question can still lead to a certain kind of progress.

3. Language is not thought, and vice versa

Let us begin to focus more closely on the relation of language and thought. We very often experience our thought as "talking to ourselves" — we actually hear snatches of phrases or sentences in our head — and it's very tempting therefore to characterize thought as some sort of inner speech. I want to argue, though, that thought cannot be simply bits of language in the head, that it is a different kind of brain phenomenon. These arguments are by no means new (see Dascal 1995 for a history of the dispute), but it's worth rehearsing them again in this context.

First, thinking is largely independent of what language one happens to think in. A French speaker or a Turkish speaker can have essentially the same thoughts as an English speaker can — they're just in French or Turkish.² The point of translation between languages is to preserve the thought behind the expression. If different languages can express the same thought, then thoughts cannot be embalmed in the form of any single language: they must be neutral as to what language they are to be expressed in. For instance, the same thought can be expressed in English, where the verb precedes the direct object, and in Japanese, where the verb follows the direct object. Hence the form of the thought must not depend on word order. *Language*, by contrast, *does* depend on word order: you (or your brain) have to choose a word order in order to say a sentence — or even just to hear it in your head. Let's think also about the experience of bilinguals who can "think in two languages". We would like to be able to say their thoughts are essentially the same, no matter which language they are "thinking in". This is possible only if the form of thought is *neither* of these languages.³

In fact, linguistic expression can exist without thought of comparable complexity. A personal example is my Bar Mitzvah-level Hebrew: I can rattle off substantial passages quite fluently — or hear them in my head if I so choose; but there are large portions of them for which I haven't the foggiest notion of what thought they express. What I have learned is not simply random noise: phonology, stress, and many elements of morphological and syntactic structure are present. It is more like knowing *Jabberwocky* ("Twas

brillig, and the slithy toves/Did gyre and gimble in the wabe”), which surely invokes knowledge of English, than it is like knowing reverse *Jabberwocky* (“Ebaw eth ni elbmig dna eryg did/Sevot ythils eth dna, gillirb sawt”), which has nothing to do with knowledge of English. Hence this situation justifiably can be said to involve language without accompanying thought.

A more robust case comes from children with Williams Syndrome, a congenital defect that results in pretty profound mental retardation (Bellugi, Wang, and Jernigan 1993). What is interesting about these individuals is that at first glance they seem actually rather precocious, because they chatter away in animated fashion, using all kinds of advanced vocabulary. But their language skills turn out to be an isolated high point in an otherwise bleak intellectual landscape; they evidently cannot use their advanced language to reason very well.

Yet another case is the linguistic idiot savant studied by Smith and Tsimpli (1995). This is an individual retarded in nearly every respect except for the learning of languages, in which he displays remarkable ability, having acquired substantial fluency in Danish, Dutch, Finnish, French, German, Greek, Hindi, Italian, Norwegian, Polish, Portuguese, Russian, Spanish, Swedish, Turkish, and Welsh. Interestingly, although his word-for-word translation and use of grammatical inflection is excellent, he is poor at making use of contextual relevance to refine his translations, suggesting some dissociation between using language and understanding it in any depth.

In each of these cases, then, linguistic expression is highly structured but is not connected to commensurately complex meaning. Conversely, complex thought can exist without linguistic expression. Think about people like Beethoven and Picasso, who obviously displayed a lot of intelligence and deep thought. But their thoughts were not expressible as bits of language — their intelligence was in the musical and visual domains respectively. (If you are tempted to respond that music and visual art are languages of a sort, I’ll agree that they’re modes of communication, but not that they’re *languages*; see Lerdahl and Jackendoff 1983 for discussion.) Particularly in Beethoven’s case, we can actually trace his thought through the evidence of his sketchbooks, which record the evolution of complex musical thoughts over periods of years, with very little if any linguistic commentary attached.

For a much more mundane example, think of the motor intelligence displayed in the simple act of washing dishes. It is well beyond current understanding in computer science to program a pair of robot eyes and hands

to wash dishes in the manner that we do, with the skill and delicacy that any of us can muster. I suspect it would be impossible to train an animal to carry out such a task, much less induce it to have the persistence to want to finish the job (this parallels in part Dennett's (1995) example of tending a fire). Yet very little of this skill is verbalizable.

Turning more directly to animals, there are ample arguments for reasoning in nonhuman primates, going back to Köhler (1927) and continuing in a rich tradition to the present. If apes and monkeys can think without language, at least to some degree, then we have to acknowledge the independence of thought from language. To make this point more vividly, I want to draw your attention to a particular phenomenon called 'redirected aggression', described by Cheney and Seyfarth (1990). From extended observation of and experimentation with vervet monkeys in the wild, Cheney and Seyfarth show that if monkey X attacks monkey Y, there is a strong likelihood that monkey Y will shortly thereafter attack some other member of X's kin-group. Consider what reasoning must be attributed to Y to explain this behavior: Y must know (1) that X attacked Y, (2) that retaliation is a desirable response, (3) that this other monkey, call her Z, is a member of X's kin-group, (4) that kin-groups count as some sort of equivalence class for attack and retribution, and (5) that facts (1)-(4) give Y reason to attack Z. This seems to me like fairly complexly structured thought, and, other than the actual attack, the facts involved are related to perception only very abstractly. (Cheney and Seyfarth mount extended arguments against regarding this behavior as any sort of simple stimulus generalization; their explicit goal is to convince animal behaviorists that reasoning is indeed taking place.)

The picture that emerges from these examples is that although language *expresses* thought, thought itself is a separate brain phenomenon.

4. Language is conscious, thought is not

Next I want to point up a difference between language and thought with respect to consciousness. Compare *rhyming* with *entailment*. When we observe that two words rhyme, say *mouse* and *house*, we can easily see why: the ends of the two words sound the same: *-ouse*. That is, the relevant parts of words in a rhyming relationship are immediately transparent to awareness. By contrast, consider an entailment relationship, for instance if *Bill killed Harry*,

then *Harry died*. This entailment is intuitively trivial. But exactly what parts of the first sentence are responsible for entailing exactly what parts of the second sentence? You might say "That's just what killing is: making someone die". But this merely restates the problem: there's nothing about the *form* of the word *kill* that makes it obvious that it has anything to do with the *form* of the word *die*. The relationship of entailment *feels* to us as obvious and mechanical as the relationship of rhyming, but at the same time we can't put our finger on what parts of the words make the entailment mechanical. Entailment involves an intuitive step, a step that is opaque to awareness and that — if we attempt to focus on it too closely — is a little magical.

In this light, we can regard the development of formal logic, from the Greeks to the present, as a series of attempts to make explicit the steps involved in reasoning, to lay bare the mechanical principles of thought. Such an enterprise has always been deemed necessary precisely because these principles are *not* evident in natural language. By contrast, until the advent of modern phonology, no one has found it necessary to uncover the mechanical principles behind rhyming: they are self-evident.

Of course, a major difference between rhyming and entailment is that rhyming is a relation between the linguistic *form* of words, while entailment is a relation between the *meanings* of sentences, between the thoughts the sentences express. In Jackendoff (1987; henceforth C&CM), I proposed that this difference is a key to understanding why the former is transparent to awareness and the latter is not. Figure 2 schematizes my analysis.

The idea is that linguistic *forms* are available to awareness — we can

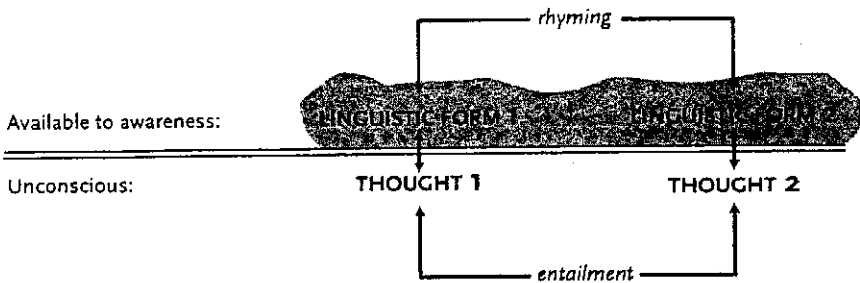


Figure 2

consciously analyze them into their parts. Thus the structures involved in judging rhymes are consciously available. On the other hand, the forms of thought, although expressed by conscious linguistic forms, are not themselves available to consciousness. Thus linguistic form 1 entails linguistic form 2 only indirectly, via a principled relationship between the thoughts they express. Since the structures involved in the relationship among thoughts are not consciously available, entailment has this intuitive, quasi-magical gap in the way it is experienced.

More generally, I am inclined to believe that thought per se is *never* conscious. This may sound radical, but consider: When we are doing what we call conscious thinking, we are usually experiencing a talking voice in the head, the so-called stream of consciousness. For most of us educated people, this voice never shuts up — we have to do Zen or something to make it quiet in there. But we have just seen that the thought is not the same as the language: the language is rather a consciously available expression of the thought. If we can catch ourselves in the act of thinking, then, it is because the linguistic images in our heads spell out some of the steps.

Remember too that there are those times when an inspiration jumps into awareness; it comes as if by magic. Of course, as good cognitive scientists, we're not allowed to say it's magic. Rather, we have to assume that the brain is going about its business of solving problems, but not making a lot of conscious noise about it; thought is taking place without being expressed as language. So when the final result pops out as a linguistic expression ("Hey! I've got it! If we just do such-and-such, everything will work fine!"), it comes as a surprise.

So far, then, I'm advocating that we become aware of thought taking place — we catch ourselves in the act of thinking — only when it manifests itself in linguistic form.

This is however an oversimplification, because there exist other conscious manifestations of thought, such as visual images. For instance, if I say *Bill killed Harry*, you may have a visual image of someone sticking a knife into someone else, that second person bleeding and falling down and ceasing to breathe, and you say "Oh yes, Harry died!". So you might suspect that the connection between the two thoughts is made in the visual image, a mental phenomenon that does appear in consciousness.

The problem with this solution was pointed out by Bishop Berkeley back

in the 18th century; the modern version of the argument appears in Fodor (1975). An image is too specific. When I said *Bill killed Harry*, your image had to have Bill stabbing or strangling or shooting or poisoning or hanging or electrocuting Harry. Any of these count as killing. And Harry had to fall down, or expire sitting down, or die hanging from a rope. Any of these count as dying. So how can any of them be *the* concept of killing or dying? That is, the thoughts expressed by the words *kill* and *die*, not to mention the connections between them, are too general, too abstract to be conveyed by a visual image.

A second problem, emphasized by Wittgenstein (1953), is that of identification. How do you know that those people in your visual image are Bill and Harry respectively? There's nothing in their appearance that gives them their names. (Even if they're wearing sweatshirts with *Bill* and *Harry* emblazoned on them, that still doesn't do the trick!)

A third problem: What is the visual image that corresponds to a question like *Who killed Roger Rabbit*? This sentence clearly expresses a comprehensible thought. But there is nothing that can be put in a visual image to show that it corresponds to a question rather than a statement, say *Someone unknown killed Roger Rabbit*. In fact, what could the image be even for *someone unknown*? And the situation is still worse when we try to conjure up a useful image for *virtue* or *social justice* or *seven hundred thirty-two*.

My view is that visual images, like linguistic images, are possible conscious manifestations of thought, but they are not thoughts either. Again, this is nothing new.

But now let us put this together with Figure 1. Figure 1 distinguishes a sort of shell of sensory brain phenomena that are completely unconscious. In the case of vision such phenomena include such things as fusion of the two retinal images, edge detection, and stabilization of the visual field in spite of eye movements. In the case of language these peripheral phenomena include at least the analysis of the auditory signal into phonemes or speech sounds.

I am now suggesting that there is as well a central core of brain phenomena — thought — that is inaccessible to consciousness. So the phenomena that *are* conscious form a sort of intermediate ring between the sensorimotor periphery and the cognitive interior. Figure 3 is a crude picture of this. (Figure 2 can be regarded as a segment of Figure 3.)

Since we act on the world as well as perceive it, Figure 3 includes not

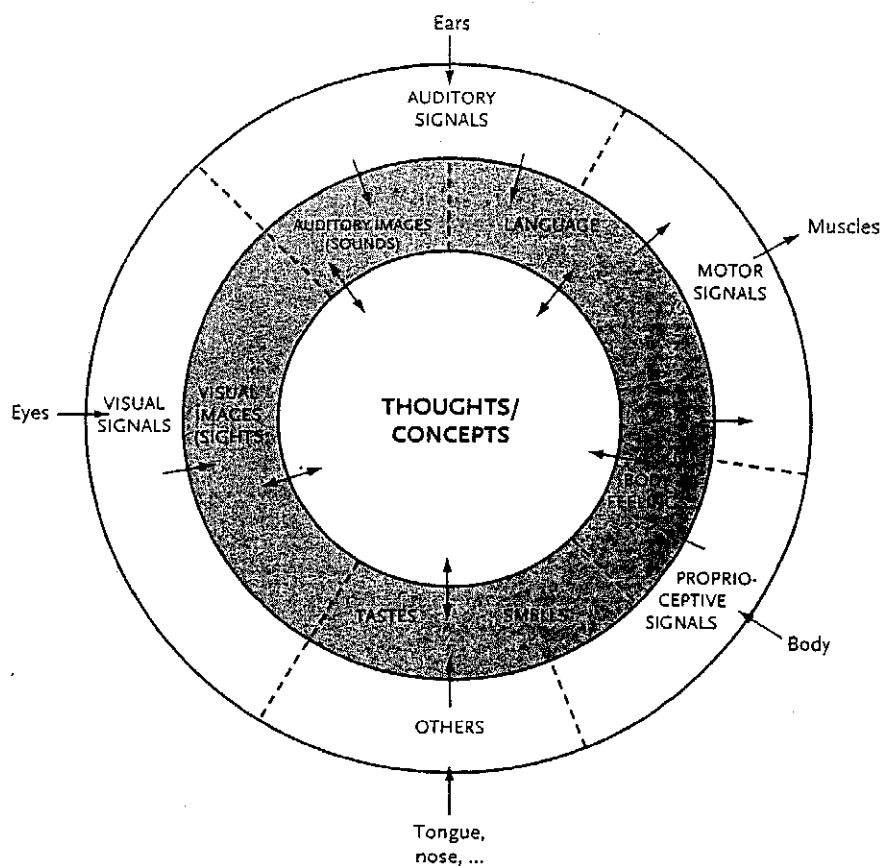


Figure 3

only sensory information coming into the brain, but also motor instructions going out. Only the phenomena in the shaded ring are conscious. The outer peripheral ring is unconscious, as in Figure 1, and the inner cognitive core is also unconscious.

Figure 3 schematizes a position I have called the Intermediate Level Theory of Consciousness, worked out in a good deal more detail in *C&CM*. With respect to language and thought, the idea is that the *form* of our

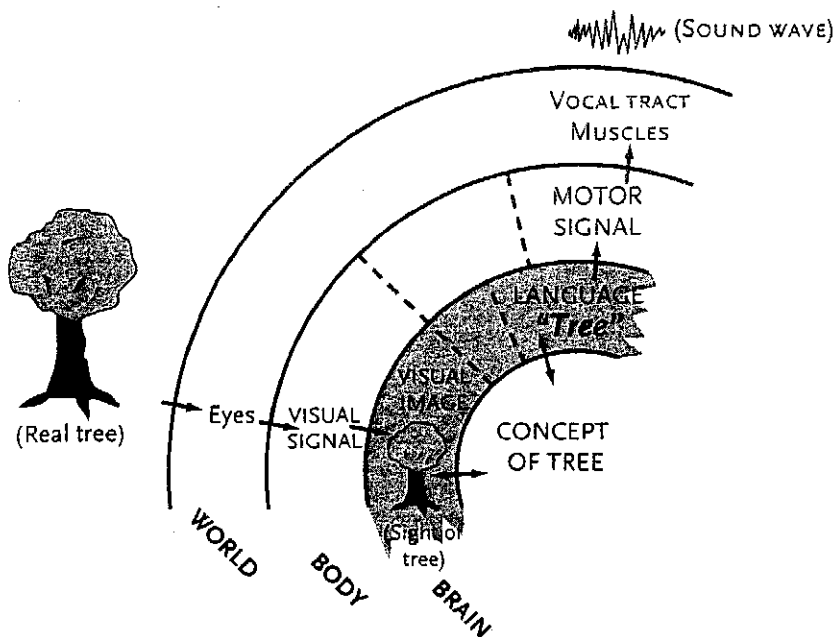


Figure 4

experience is driven by the form of language, especially by phonological structure. We experience organized sounds. On the other hand, the *content* of our experience, our understanding of the sounds, is a different organization — a different brain software, if you like — called conceptual structure.⁴ The organization of this content is completely unconscious.

Thought often drives language: the presence of a conceptual structure in the brain often causes the brain to develop a corresponding linguistic structure, which we may either pronounce or just experience as linguistic imagery. Conversely, linguistic structures in the brain (most often created in response to an incoming speech signal) invariably drive the brain to try to create a corresponding thought (the meaning of the heard utterance). Consequently, many of our thoughts have a conscious accompaniment: the linguistic structures that express them.

Thought can be driven by other modalities as well. For instance, consider Figure 4.

In response to viewing a tree in the world, the visual system constructs a

representation of the form of the tree — and it also drives the conceptual system to retrieve the *concept* of a tree. That is, our understanding of what we see is not a consequence of visual images alone, but also of the conceptual organization connected to those images. As in the case of language, we don't have direct awareness of our understanding of trees; we just have a perceptual awareness of the visual form of what we're seeing. The visual form is in the intermediate shell of brain processes in Figure 3, but the understanding is in the unconscious central core.

In turn, the unconscious central core links the visual perception with language. If the visual system drives the conceptual system to develop a conceptualization of the tree, then the conceptual system can drive the language system to create an utterance such as "Hey, a tree!". Thus, in Figure 4, the two conscious parts, the visual image and the word, are linked together by the unconscious concept of a tree.

Part of the idea behind Figure 3 is that the same general architecture is preserved in other organisms, at least those with higher mammalian brain structure (though not necessarily reptiles or Martians). For instance, a monkey will have essentially the same linkages among representations as we do, lacking only the link provided by language. In fact, an organism might have other modalities of experience that humans don't have. Bats presumably have some kind of experience through their echolocation modality whose quality we can't imagine — *hearing* shape and distance and motion! This follows from the Intermediate Level Theory, which says that somewhere between sensation and cognition lies a level of brain phenomena that is conscious. So we can imagine taking the basic model of Figure 3 and adding and subtracting different modalities of experience to imagine the forms of experience for different organisms.

The overall spirit of this analysis thus fits into the view that the mind is modular in structure — that it has a lot of specialized parts or modules, each of which is good at doing one thing and doesn't know much about what the other parts are doing. In Figure 3, the different sectors of each ring stand for different modules.

What is missing in the best-known exposition of modularity, that of Fodor (1983), is an account of how the modules interact with each other. In Fodor's approach, the language perception faculty, for instance, derives linguistic representations from auditory signals, but Fodor does not specify how these linguistic representations get converted into thoughts. My view is that,

almost as a point of logical necessity, we need also modules that interface between faculties: a module that dumbly, obsessively converts thoughts into linguistic form and vice versa, another that dumbly, obsessively converts visual images into thoughts and vice versa, and so forth. Of course, not all thoughts lend themselves to visual imagery or linguistic form, so the interface modules just do the best they can. The places where they fail are just those places where we have found gaps between the conscious forms of linguistic and visual experience on one hand and the needs of thought and reasoning on the other.

5. The significance of consciousness again

Let us continue to look at Figure 3, and ask what it says about consciousness. Section 2 mentioned that many people want to ascribe to consciousness some cosmic significance — they want to say our consciousness is what makes us human, it's connected to our highest capacities, it's what separates us from the animals. If anything like Figure 3 is correct, the significance of consciousness can't be anything like that. If there is anything that makes us human, that has enabled us to build great civilizations, it is our capacity for thought — *which isn't conscious!* Consciousness is not directing our behavior. Rather, it is just a consequence of some intermediate phenomena along the chain of connection between perception (the periphery) and thought (the core).

This goes strongly against intuition, of course. But it stands to reason that it should. The conscious phenomena are all we can know directly about our minds. As far as consciousness is concerned, the rest of the mind simply doesn't exist, except as a source of miracles that we get so accustomed to that we call them 'common sense'. Hence we naturally ascribe any intellectual powers we have to consciousness. I submit that, if we really look at the department of so-called miracles, we find that most of the interesting work of intelligence is being done in that department — not in the conscious department. (My impression is that this has certainly been one of the lessons learned from AI.)

Notice that this conclusion does not demean our humanity in any way. We can still do whatever we can do. Maybe it does reduce the chance that it would be interesting to have a consciousness that continues beyond our death — an immortal soul. But, after all, wishing for something doesn't make it true.

We'll have to look for reasons other than the hope of immortality to find our lives meaningful.

This completes the first part of my assignment, in which I've tried to dissociate language from thought and to show that they are distinct brain phenomena. According to this picture, the mere fact of having language neither makes thought possible, nor does it have a direct effect on thought processes. In the rest of this paper I want to suggest three ways in which language has important *indirect* effects on thought processes.

6. First way language helps us think: Linguistic communication

The first way language helps us think is fairly obvious. By virtue of having language, it is possible to communicate thought in a way impossible to nonlinguistic organisms. Language permits us to have history and law and science and gossip. The range of things we can think about is not limited to what we can dream up ourselves — we have all of past and present culture to draw upon. That is, language permits a major enhancement in the *range* of things the thought processes can pertain to, even if the thought processes may be exactly the same.

Without language, one may have abstract thoughts, but one has no way to communicate them (beyond a few stereotyped gestures such as head-shaking for affirmation and the like). Without language, there is no way to communicate the time reference of events: does a pantomimed action describe a past action or a desired future action, for instance? Without language, one may have thoughts of being uncertain whether such-and-such is the case, but one cannot frame questions, which communicate one's desire to be relieved of that uncertainty. And so on and so on. As a result of having language, vastly more of our knowledge is collective and cumulative than that of nonlinguistic organisms. To be sure, other higher mammals have societies and even perhaps proto-cultures (Wrangham et al. 1994). But the collectivity and cumulativity of human thought, via linguistic expression, almost certainly has to be a major factor in the vast expansion of knowledge and culture in human societies. Good ideas can be passed on much more efficiently.

That much, I should think, is essentially unquestionable. I can imagine pushing this line a little further, though. If the possession of a language capacity makes it possible to have much richer collective and cumulative knowledge, then it can conceivably create evolutionary pressure towards

thought processes better adapted to deal with this abundance — perhaps greater speed or efficiency and greater short-term memory, to facilitate on-line thought, or greater long-term memory to facilitate accumulating more knowledge altogether. Creatures better adapted to make cognitive use of the advantages proffered by linguistic communication would presumably have selective advantage over those less adapted.

Whether such pressure has *actually* changed the inherent reasoning capacity of humans is an open question. We do not know when language evolved nor whether it appeared all at once or (as advocated by Bickerton 1990) in stages. Still less do we know about the cognitive capacities of prehistoric humans, except to the extent that we find cultural artifacts such as tools, ornaments, and habitations. I don't think we yet have any way of figuring out how much less "intelligence" than modern humans one could have and still develop prehistoric cultures. Thus, given that the timing of the evolution of both language and enhanced intelligence are unknown, it seems pretty useless to speculate which pushed which in the course of evolution. But it is a *potential* way that language could have helped us develop the capacity to think, over the course of human evolution.

7. Second way language helps us think: Making thought available for attention

The enhancement of thought by virtue of linguistic communication might strike many readers as a sufficient connection between language and thought. But I want to dig a little deeper, and explore the relation between language, thought, and consciousness that emerged in the Intermediate Level Theory sketched in Figure 3.

7.1 Introduction

Imagine what it would be like to be an organism just like a human except lacking language. According to Figure 3, you could see and hear and feel exactly the same sensations, perform the same actions, and in particular think exactly the same thoughts. There would be two differences. First, the one just discussed: You couldn't benefit from linguistic communication, nor could you make your knowledge available to others through the linguistic medium. So the range of things you could think about would be relatively impover-

ished. But I want to concentrate on another difference, the phenomenological difference: you couldn't *experience* your thought as linguistic images. That is, a whole modality of experience would be simply absent — a modality which, as pointed out earlier, is very important to human experience.

What would the effect be? Let's return to the monkeys who practice redirected aggression. Recall what happens: Monkey X attacks monkey Y, and then monkey Y attacks not monkey X, but monkey Z, who happens to be a member of X's kin-group. (It's not always that bad: sometimes Y will go and be nice to Z instead, as if to indirectly make up with X or appease X.) Recall that something like the following pieces of reasoning are logically necessary to account for this behavior:

- (1) a. X attacked me.
- b. An attack on me makes me want to take retaliatory action (alternatively, to make up).
- c. The members of an attacker's kin-group are legitimate targets for retaliation (or making up).
- d. Z is a member of X's kin-group.
- e. Therefore, I will attack Z.

As we noticed earlier, factors b, c, and d are abstract, that is, not connected directly to perception in any way — factor c especially so. Now what seems so intuitively weird about this chain of reasoning is that we can't imagine a monkey saying things like (1a-e) to herself. And indeed she doesn't, because she doesn't have language. Yet thoughts very much like those expressed by (1a-e) *must* be chained together in the monkey's head in order to account for the behavior.

How do we escape this conundrum? Given what we've done up to this point, the answer is fairly clear: the monkey has the thoughts, but she doesn't hear the corresponding sentences in her head, because she has no linguistic medium in which to express them. Consequently, the monkey doesn't *experience* herself as reasoning through the chain in (1). All she experiences is the outcome of the reasoning, namely an urge to attack monkey Z, and maybe an image of herself attacking monkey Z. The monkey's experience might be like our sudden urges, for no apparent reason, to go get a beer from the fridge. A human, by contrast, does have the linguistic medium, and so, when a human has thoughts like those expressed in (1), they may well generate experienced thought, namely a conscious plotting of revenge.⁵

The point is this: Only by having a linguistic modality is it possible to experience the steps of any sort of abstract thought. For example, one can't directly *see* that monkey Z is a member of X's kin-group — this is an abstract predicate-argument relationship. But this relationship can be made explicit in consciousness by using the words "Z is kin to X". Similarly, the notion of retaliation — performing some action *for a particular reason* — is unavailable to nonlinguistic consciousness. A reason *as a reason* cannot be represented in a visual image. But by using the word *because* to relate two propositions, the linguistic modality can make reasons as such available in consciousness.

Still, if you have been following the argument up to here, you may well say, "So what?". The fact that these thoughts appear in experience doesn't change them, and it doesn't change the steps of reasoning that can be performed on them. Language is just a vehicle for externalizing thoughts, it isn't the thoughts themselves. So having language doesn't enhance *thought*, it only enhances the *experience* of the thought. Maybe thinking is more fun if we can experience it, but this doesn't make the thought more powerful. And that is about where I left matters in C&CM.

7.2. *The relationship of consciousness and attention*

However, I now think this is not the end of the story. I want to suggest that, by virtue of being available to consciousness, language allows us to *pay attention* to thought, yielding significant benefits. To see this, let us set language aside for a moment, and think about the relation of consciousness and attention.

It is often pointed out that consciousness and attention are intimately interlinked. An often-cited example involves driving along having a conversation with a passenger. On the standard account, you are said to be unconscious of the road and the vehicles around you — you navigate "without awareness" — until your attention is drawn by some complicated situation, perhaps a lot of traffic weaving through the intersection up ahead; and you have to suspend conversation momentarily while attending to the traffic. From anecdotal experiences like this, many people have concluded that consciousness is a sort of executive (or traffic cop) that resolves complicated situations in the processing of information, and that we are only conscious of things that are hard to process. (I understand Minsky 1986 as taking essentially this position, for instance.)

Notice that such a view goes well with the idea that consciousness is a top-level mental faculty, deeply entwined with our higher intelligence, free will, and so forth. In working out the Intermediate Level Theory of consciousness (Figure 3), I have challenged that overall prejudice, and so I am automatically suspicious of drawing such a connection between consciousness and attention.

Let us therefore look a little more closely at the situation of driving and conversing at the same time. Are you really *unconscious* of your driving? You may not remember what you did afterward, and you're certainly not paying serious attention to it, but at the same time you're not completely unaware of what you're doing, in the way, say, that you're unaware of your saccadic eye movements or of the flow of serotonin in your brain, or for that matter of what's going on in your environment when you're asleep. Rather, I think it's fair to say that when you're driving and conversing, you're at least vaguely aware of the flow of traffic. It's not as if you're blind.

Consider also another situation. Suppose you are just lying on a beach, idly listening to the waves, wiggling your toes in the sand, watching people go by. There's no sense of anything being hard to process, it's just utter relaxation. But surely you're conscious of all of it. It seems as if the people who want to view consciousness as a high-level decision-making part of the mind tend to forget about such cases.

We certainly need to make a distinction between fully attentive awareness and vague awareness. But *both* are states of consciousness — something is in the field of consciousness. The goings-on that you're only vaguely aware of may be less vivid or immediate, but they are still "there" for you. To account properly for such phenomenology, we must make our terminology a bit more precise than common language here, because we often *do* say "I wasn't conscious of such-and-such" when we mean something like "Such-and-such didn't attract my attention".

My sense is that consciousness has nothing at all to do with processing difficulty or executive control of processing: that's the function of *attention*. *Attention*, not consciousness, is attracted to those parts of the perceptual field that potentially present processing difficulty — sudden movements, changes in sound, sudden body sensations, and so forth.

So we are led to ask another of those superficially obvious questions: What happens in the brain when one pays attention to something? Most research on attention seems to have focused on what it takes to *attract*

attention, or on the maximal capacity of attention (see Allport 1989 for a survey). I am asking a different question: What good does attention do once it is focused on something?

A traditional approach to attention (Broadbent 1958, for example), takes the position that the function of attention is to filter out incoming information that would "overwhelm" the processing capacity of the brain. On this view, the function of attention is to keep out everything not attended to. I would prefer a more positive characterization of the same intuition: the processing capacity of the brain for dealing with incoming signals is indeed limited, but resources can be distributed in different ways. If resources are distributed more or less evenly, the result is a more or less uniform degree of detail throughout the perceptual field. Alternatively, resources can be distributed unevenly, so as to enhance certain regions, but at the price of degrading others. I take attention to be this function of selective enhancement.

There seem to be at least three things that we can do with a percept when we pay attention to it, things that we cannot do with nonattended material. First, as just posited, focusing attention on something brings more processing resources to bear on it, so that it can be resolved faster and/or in greater detail. In turn, this extra detail is what makes the consciousness more vivid and immediate. Because the remaining phenomena get fewer processing resources, they are not resolved in as much detail, and so they are vaguer in consciousness. On the other hand, if attention is not particularly focused, say, when lying on the beach, the quality of consciousness is perhaps more uniform across the perceptual field.

A second thing that happens when we pay attention to an object is that we can "anchor" it — stabilize it in working memory while comparing it to other objects in the environment, shifting rapidly back and forth, or while retrieving material from memory to compare with it. We can also anchor the percept while "zooming in" on details and attending to *them*. Another aspect of anchoring is the ability to track moving objects against a background, if we pay attention to them — or, more difficult, to track one moving object amidst a swarm of similar objects (Culham and Cavanagh 1994).

A third thing that happens when we pay attention to something is that we can individuate it and remember it as an entity in its own right. If we take in a particular visual figure just as part of the texture of the wallpaper, for example, we will never remember it. But once we pay attention to it, we can go back and pick out *that very one* — or notice that it isn't there if it has moved away.

Now I have to be careful about something. I have been saying sometimes that attention is drawn by an object in the world, sometimes that it is drawn by a percept in the head. Which is right? If we want to be precise, we have to say a percept. Attention is a process going on in the brain — so it *cannot* be directed by things moving out there in the external world. Attention is not a little person in the brain watching the world and pointing a spotlight at interesting things out there so they can be seen better. Rather, attention has to be directed by the character of brain phenomena that have occurred in response to things in the external world of potential interest to the organism — a crucial difference. It all has to happen inside the brain.

This now raises an interesting question, which I think has not been asked in exactly this way before: Of all the brain phenomena that may take place in response to, say, a sudden motion or noise in the world, which ones are capable of being focused on by attentional processes? If we consider the phenomenology of attention, an answer suggests itself: *We can pay attention only to something that we are conscious of.* We may not *understand* what we are paying attention to (in fact that may be the reason we are paying attention to it) — but we must certainly be *aware* of it. In other words, the representations that fall into the intermediate ring of Figure 3 play some necessary role in the functioning of attention; perhaps we can think of these representations as being potential “handles” by which attention “grasps” and “holds onto” percepts.

This puts a new twist on the more or less standard story that consciousness drives attention. Instead of thinking of consciousness as the smart (or even miraculous) part of the mind that determines which percepts need attention, I am claiming that consciousness happens to provide the basis for *attention* to pick out what might be interesting and thereby put high-power processing to work on it. In turn, the high-power processing resulting from attention is what does the intelligent work; and at the same time, as a byproduct, it enhances the resolution and vividness of the attended part of the conscious field.⁶

7.3. *Language provides a way to pay attention to thought*

Now let us go back to language and see why I took this detour on consciousness and attention.

Remember two points: first, language provides a modality of consciousness that other animals lack; second, language is the only modality that can

present to consciousness abstract parts of thought like kinship relations, reasons, hypothetical situations, and the notion of inference. Only through language can such concepts form part of *experience* rather than just being the source of intuitive urges.

At the point we left off in section 7.1, we could see that although it might be more fun to experience one's reasoning through language, it would not yet help reasoning in any way. What we just said about attention, though, adds a new factor: having linguistic expressions in consciousness allows us to pay attention to them. And now the extra processing power of attention can be brought into play.

Consider yet again the monkey who is calculating out some redirected aggression. Let's make it instead a person, say Bill, who has a language module and with it linguistic awareness. Instead of just experiencing an urge to attack whoever it is, say Joe, Bill's concepts may drive his language faculty to produce some utterance (he may actually say it, or he may just hear it in his head):

(2) I'm gonna KILL that guy!

First of all, notice that the language faculty forces Bill to package his thought in some particular way, in accordance with words that happen to exist in English. This feeds back into the thought and refines it: Bill's generalized sense of aggression gets molded into a threat to murder Joe, as opposed to, say, just insult him.

Now, if Bill doesn't pay attention to his utterance, it's just mumbling and has no further effect — it's just a simple statement of generalized aggression, a bit of texture in the auditory wallpaper. But suppose Bill does listen to himself — suppose he pays attention to his utterance. Then the sentence gets anchored, more computing power is devoted to it, and details can develop:

(3) KILL? Or just mangle, or just insult?
 That guy? Or someone else, or his whole family?
 Gonna? When? Tomorrow? Next week?
 Why? What do I hope to accomplish? Is there a better way?
 How? Should I shoot him? Stab him? Poison him? Club him?
 What then? What will happen if I kill him?

And attention to some of these details may lead to elaboration of further details:

(4) Club him? With what? A baseball bat? A fence post? ...

And any promising options may be remembered. Notice again that this is all from attention being paid to the *sentence*. Without the sentence as a conscious manifestation of the thought, attention could not be applied, since attention requires some conscious manifestation as a "handle".

We should also notice hidden in these ruminations of Bill's the expression of some concepts that would not be available without language, for example the notion 'next week'. What does it take to think about weeks? A nonlinguistic organism can obviously detect patterns of light and darkness, and respond to diurnal patterns. But it takes the word *day* to abstract this pattern out as a sort of object, so that attention can be drawn to it as a constancy. The word is a perceptual object that anchors attention on the pattern and allows it to be stored as a repeatable and retrievable unit in memory.

What about a *week* — a unit of seven days? This is a completely nonperceptual unit. I don't think it could be conceived of at all without linguistic anchoring. Even if such a unit were potentially available as a concept, it couldn't be accessed without having language to hold it in attention, which enables us to stow such a unit away for future use. I think it fair to say that although nonlinguistic organisms may be able to develop a concept of a day, they will never attain the concept of a week.

To sum up this overly long section: Language is the only modality of consciousness in which the abstract and relational elements of thought are available as separable units. By becoming conscious, these elements of thought become available for attention. Attention in turn refines them, both by anchoring and drawing out details, and also by concretizing or reifying conceptual units that have no stable perceptual base.

8. Third way language helps us think: Valuation of conscious percepts

The third way language helps us think is related to the second. To understand it, we have to step back yet again and examine another property of consciously available percepts.

What's the difference between the appearance of something that looks familiar and that of something that doesn't look familiar? In general, nothing: as you get used to the appearance of some novel picture, let's say, the

appearance doesn't change — it just somehow feels different than when it was novel. The same for the sound of a new tune as you gradually get to know it. Or, for the classic case, suppose someone is running a psychological experiment on you and asks you which nonsense syllables are the same as the ones they gave you yesterday. What's the difference between the familiar ones and the novel ones? They all sound more or less like nonsense, except that some come with this feeling of familiarity and some don't.

I'm going to call this feeling of familiarity or novelty attached to a percept a *valuation* of the percept. We can even have this sense of familiarity without clearly knowing what we're seeing — "I'm sure I know you from somewhere, but I'm damned if I can remember who you are or when we met". When we have a *déjà vu* experience, we somehow get a feeling of familiarity attached to an object or situation that we rationally know we've never experienced before — that is, *déjà vu* is an error or illusion of valuation.

The familiar/novel distinction is not the only valuation that the brain applies to percepts. Another one is the distinction between percepts that are taken to be imaginary and those that are taken to be real. ("Is this a dagger I see before me?") This valuation is a bit trickier to isolate, because things we judge to be real tend to be vivid and substantial in consciousness, whereas things we judge to be imaginary tend to be fuzzy and fleeting. But in certain limiting cases, it is possible to see that the valuation can make an independent difference. Suppose you are trying to make your way through a thick fog, and you are just not sure what you're seeing. Was that blurry motion you detected something really out there, or do you just think it is? Was that noise you heard real, or just your imagination? When for some reason perception is impeded, the very same percept may come to be judged either real or imaginary, and there is nothing clear about its appearance that helps us make the decision.

Another sort of limiting case is dreams, where things may seem quite real that upon awakening are judged imaginary. "Don't be frightened: it was only a dream!" — we say. Like *déjà vu*, this is an error or illusion of valuation.

A related valuation concerns whether a percept is externally or internally initiated. What do I mean by this? Consider a visual image you may get as a result of my telling you "Imagine a pink elephant", and compare it to a similar image you might get as a result of drinking too much, where it "just pops into your head". The former image, you feel, is under your control; the latter is not. Yet both are in fact generated by your own brain activity, and you can't catch yourself in the act of making either image happen. There's just this mysterious, miraculous "act of will" — or its absence.

Or think of the sense of blinking your eyes voluntarily as opposed to doing it automatically. The movements themselves are essentially the same, and both ultimately require similar nerve activation of the muscles. But they feel different, namely in or out of your control.

Of course the hallucinations of schizophrenics are errors in both of these latter valuations: they hear voices that they take to be real and externally generated ("God is really speaking to me", they say), whereas the voices are in fact imaginary and internally generated.

The general idea, then, is that our cognitive repertoire contains a family of valuations, each of which is a binary opposition that can determine part of the "feel" of conscious percepts. These three and a number of others are discussed in C&CM, chapter 15 (where they were called "affects", a term I was never quite comfortable with). Valuations of percepts have not to my knowledge been singled out for attention elsewhere in the literature. What is curious about them is that they are not part of the *form* of consciousness; as stressed above, they're more like a feeling associated with the form.

But — if we have language, we *can* give these feelings some form: we have words like *familiar*, *novel*, *real*, *imaginary*, *self-controlled*, *hallucination* that express valuations and therefore give us a conscious link to them. This conscious link permits us to attend to valuations and subject *them* to scrutiny: Is this percept really familiar, or is it a *déjà vu*? Is it real or a dream? and so forth. A dog awakening from a dream may be disturbed about what happened to the rabbit it was chasing; but it cannot formulate the explanation "It was a dream". Rather, something else attracts its attention, and life goes on. But with language, we can fix on this valuation as an independent object in its own right, and thereby explain the experience — as well as recognize a category of experiences called "dreams".

The plot thickens. Because linguistic forms are percepts too, they can themselves be subject to valuation. For instance, what's going on when you judge that some sentence is *true*? There is nothing about the literal sound of a true sentence that's different from the literal sound of a false sentence, yet we say "it sounds true to me". That is, the sense of a sentence being true or false is also — from a psychological point of view — a kind of valuation. It is altogether parallel to a judgment that a visual percept is something really out there. Similarly, the concept that we express by *suppose that* or *if* is a valuation that suspends judgment, parallel to evaluating some visual image as imaginary and internally produced.

But now let us combine this with the previous point. Like other valuations, the valuations of language can be expressed in language, with words like *true*, *not*, *if*, and so forth. Therefore *these* valuations can be attended to as independent objects in their own right and focused on, and so we get full-blown recursion — a thought about the valuation of another thought, the larger thought having its own valuation. We can express thoughts like “Suppose I am incorrect about such-and-such ... then such-and-such other belief of mine is also false”. That is, it is precisely by virtue of thoughts having linguistic form that it is possible to reason about reasoning. There is no other modality in which valuations can be given palpable form so they can be attended to and thought about. And certainly a crucial source of the power of our reasoning is its ability to examine itself.

I don't see any evidence that nonlinguistic organisms can engage in such metareasoning. Apes and dolphins can be very clever in solving certain kinds of problems, and they can be uncertain about how to solve a problem, but I don't think they can *wonder why they are uncertain*. They may be able to believe something,⁷ but they can't wonder why they believe it, and thereby be motivated to search for evidence. It takes language to do that.⁸

9. Summing up

The first half of this paper established that language is not itself the form of thought, and that thought is totally unconscious. However, thought is given a conscious manifestation through the linguistic expressions that it drives. The second half of the paper suggested three ways in which having language enhances the power of thought:

1. Because language allows thought to be communicated, it permits the accumulation of collective knowledge. Good ideas don't get lost. This conclusion is certainly nothing new.
2. Language is the only modality of consciousness that makes perceptible the relational (or predicational) form of thought and the abstract elements of thought. Through these elements being present as isolable entities in consciousness, they can serve as the focus of attention, which permits higher-power processing, anchoring, and, perhaps most important, retrievable storage of these otherwise nonperceptible elements.

3. Language is the only modality of consciousness that brings valuations of percepts to awareness as independent elements, permitting them to be focused on and questioned. Moreover, since linguistic expressions and their valuations are also percepts, having language makes it possible to construct thoughts about thought, otherwise unframeable.

Although these conclusions may seem in the end intuitively obvious, I've tried to find my way more carefully to them, in the course of which I've challenged some fairly standard preconceptions about the nature of consciousness. The interest of the argument lies, I think, in the intricacies of the connections among language, thought, consciousness, and attention.

10. The illusion that language *is* thought

One nice thing that emerges from the present analysis is an explanation for the common-sense identification of thought with language. As pointed out in section 5, all we can know directly of our own minds are those brain phenomena that are conscious; in the terms of section 7, these are the only ones we can pay attention to. Consequently, these are the phenomena to which we ascribe responsibility for our behavior. Since the linguistic forms accompanying thought are conscious and the thoughts themselves are not, it is altogether natural to think that the linguistic form *is* the thought. Consequently, language is quite naturally taken to be the substrate for the act of reasoning. This illusion that language is thought has been the source of endless philosophical dispute (Dascal 1995). We now can see why the illusion is so intuitively persuasive.

Recognizing this illusion allows us to examine the dark side of our initial question: why language can be a *less* effective tool for reasoning than we are often prone to assume. There are at least five sorts of gaps where language does not adequately express the structure of thought. In each case, illusions develop in reasoning because language is all we have to pay attention to.

1. The smallest unit of thought that can be expressed as an independent percept is a word. Because a word is a constant percept in our experience, we treat the thought it expresses as a constant thought — even though in fact we bend and stretch the concepts expressed by words every which way. Just within the ambit of this paper, consider how the word *unconscious* in common

usage means anything from being out cold to being vaguely aware (but not noticing) — not to mention the technical use of the term I've applied here to particular brain processes. It takes careful analysis to notice the disparity among these usages, and when we're done we don't know whether the word expresses one bendable concept or a family of more rigid related ones. Intuition is not much of a guide.

This issue arises not only with complicated words like *unconscious*, but even with simple obvious words. For instance, a small industry has developed in the linguistic community studying the semantics of prepositions. People spend considerable effort trying to determine whether the preposition *in* expresses the same concept in *the coffee in the cup* and *the crack in the cup* (I think so), and whether the preposition *into* expresses the same concept in *jump into the pool* and *crash into the wall* (I think not). Whatever the correct answer turns out to be, the nature of the problem is clear: in both cases, the use of the identical word invites us to presume we are dealing with the identical concept. Yet closer examination brings to light the unreliability of such presumptions (see Herskovits 1986, Vandeloise 1991 and Jackendoff 1996 for examples).

2. The opposite side of this problem is the delegitimation of concepts for which no sufficiently precise word exists. A prime example arises with the concepts of reasoning and belief (see notes 5 and 7). If one insists that a belief is propositional, that reasoning involves relations among propositions, and that propositions are linguistic (thereby at least partly succumbing to the illusion that language is thought), then there is no term available in the language for how animals' minds organize their perception and memory and create novel behavior on the basis of this organization. One is not *allowed* to say they have beliefs and reasoning. The forced move is to attribute to them abilities for which there *are* words, for example "instinct" or "associative learning", often prefixed by "mere". The effect is to inhibit examination of what mental ability animals actually have, because there happens to be a gap in our vocabulary just where the interesting possibilities lie.

3. Not only do we fail to recognize gaps, we tend to treat all existing words as though they have references in the real world, along the lines of concrete words like *dog* and *chair*. This tendency means we're always reifying abstract terms like Truth and Language, and constructing theories of their Platonic existence — or spending a lot of effort arguing, through careful linguistic analysis, against their reification.

4. As pointed out by Lewis Carroll (1895) as well as by Wittgenstein (1953), we don't *really* know how we ultimately get from one step in reasoning to the next. How do we know that if A implies B and B implies C, then A implies C? And how do we know that any particular chain of reasoning is an example of this rule? At the bottom, we always have to fall back on this feeling of conviction, which can't be justified through any more general laws. That is, sooner or later we hit a stage of pure valuation with no language to make it conscious. Yet we think we are reasoning completely "rationally", i.e., explicitly. Worse, we often get this feeling of conviction when it's utterly unwarranted; we are prone to delude ourselves and yet feel perfectly justified. Just as in any other application of attention, the default situation is that we don't pay attention to our valuations unless we have to, unless it causes trouble.

5. As a refinement of the previous point, our sense of conviction is too often driven by our desires, by what we *want* to have come out true. We are less prone to question our convictions if they lead to desired conclusions. At the same time, such gaps in reasoning all seem perfectly rational, because they are supported — to a great enough extent — by language.

None of these gaps would be possible if language really were the form of thought. If it were, all reasoning would be completely up front, and there would be no room for weaseling around. Behind much of the development of formal logic lies the desire to provide a more satisfactory form of language, in which all terms are precise and context-free, and in which the steps of reasoning are entirely explicit — that is, in which the idealizations in points 1-5 above are not illusory but *true*.

By contrast, in the present perspective, in which language is only an imperfect expression of thought and furthermore is the only form in which many important elements of thought are available for conscious attention, these illusions are just what we would expect. And they are in large part irremediable precisely because of the architecture of the system — because of the way the interaction of language, thought, consciousness, and attention happened to evolve in our species. At the same time, as flawed as the system is from such an ideal point of view, it's all we've got, and we might as well enjoy it. There is no question that it has given our species a tremendous boost in its ability to dominate the rest of the environment, for better or for worse.⁹

Notes

1. I thus regard the computational approach as a perspective for understanding, rather than as some ultimate truth. Regarding it this way undercuts the criticisms of Searle and Edelman. For consider: Were one to take Searle's and Edelman's arguments one step further, one might legitimately claim that indeed there aren't neurons in the head any more than there are computations: actually, there are only quarks and leptons, and consequently brain function must be explained only in terms of elementary particles. Dennett (1995) has called this absurd sort of argument "greedy reductionism": it demands too much of theoretical reduction and thereby prevents anyone from understanding larger scales of organization. I submit that neuroscience, like the computational theory of mind, is just another perspective, and that it is at least premature, if not altogether illegitimate, to expect the latter to be replaced entirely by the former.
2. I say "essentially" here in order to hedge on possible "Whorfian" effects. There are undoubtedly expressive differences among languages in vocabulary, as well as in grammatically necessary elements such as tense-aspect systems and markers of social status (e.g., French *tu* vs. *vous*). Recent research by Levinson (1996) has further uncovered crosslinguistic differences in prepositional systems which affect the expression of spatial relations. But such differences must not be blown out of proportion; they are decidedly second- or third-order effects (Pullum 1991). They may create difficulties for literary translation, where style and associative richness are at stake, but no one seriously questions the possibility of effectively translating newspapers and the like.
3. At this level of generality, I concur with Fodor's (1975) Language of Thought Hypothesis. However, in other respects I disagree with Fodor; see Jackendoff 1992, especially chapters 2 and 8; see also the next section on the related issue of modularity.
4. Although I have represented "thought" in Figure 4 as a unified set of phenomena, I actually think it can be differentiated into a number of semi-autonomous components, including at least an algebraic ("propositional") mode of representation called Conceptual Structure, and a geometric ("analogue") mode called Spatial Representation; see Jackendoff (1987, 1996); Landau and Jackendoff (1993). This differentiation goes against Fodor's (1983) claim that thought is homogeneous and nonmodular. However, for present purposes, the distinctions among these components play no important role, so for the sake of exposition I have simply used the term conceptual structure to encompass them all.
5. A terminological remark: Some readers may be uneasy calling what the monkey does "reasoning", precisely because it has no linguistic accompaniment. My question then is what term you would prefer. As Cheney and Seyfarth (1990) argue at length, no notion of "mere" stimulus generalization is going to account for the behavior. Rather, whatever unconscious process is responsible must logically connect pieces of information that are parallel to those expressed in (1a-e). If you prefer to reserve the term "reasoning" for a mental process that necessarily involves linguistic accompaniment, then call what the monkeys do whatever else you like, say "unconscious knowledge manipulation". In that case, my claim is that "reasoning" in your sense amounts to unconscious knowledge manipulation whose steps are expressed linguistically.
6. Block (1995) makes a distinction between what he calls Phenomenal Consciousness (or P-Consciousness) and Access Consciousness (or A-Consciousness). In effect, I am

arguing here that the effects Block ascribes to A-Consciousness should be actually attributed to attention. If I read him correctly, what he has called P-consciousness corresponds fairly well to what I have called simply consciousness.

7. Similar remarks apply to the term "belief" here as to "reasoning" in note 5.
8. At one presentation of this paper, a member of the audience pointed out that my argument may pertain not only to animals, but also to those unfortunate individuals among the deaf who have not been exposed to sign language. According to this story, such individuals could certainly think (and, by virtue of their larger brains, better than animals). However, they would not possess the advantages in thought conferred by language. I think I have to admit this conclusion as a distinct possibility, even if politically incorrect. If ethically possible, it deserves experimentation. If *empirically* correct, it is just one more reason to make sure the deaf are exposed to sign language as early in life as possible.
9. This paper was originally inspired in reaction to a manuscript by Derek Bickerton on the relation of language and thought, which was explicit enough for me to be able to figure out where I disagreed. The manuscript has since been published as *Language and Human Behavior* (Seattle, University of Washington Press, 1995), but not in time for me to include a proper dialogue with Bickerton in the present paper. Serious thought about this material was instigated by invitations from Ohio State University (Peter Culicover) and the Goethe-Universität of Frankfurt (Gunther Grewendorf) to speak on a topic of general interest. Marcelo Dascal encouraged me to massage the talk into publishable form, despite my doubts about it. Dascal, along with Daniel Dennett, Marcel Kinsbourne, and Marjolein Groefsema, offered important comments on an earlier version. Discussion with Patrick Cavanagh helped me refine my sketch of attention. Needless to say, none of these colleagues is responsible for my folly. In addition to Ohio State and Frankfurt, I have benefited from the opportunity to present this material in talks at Dartmouth College, the Summer Institute for Cognitive Science at SUNY Buffalo, Carleton University, Vassar College, and the 1995 Language Acquisition Research Symposium in Utrecht; I wish to thank those who invited me and those who had to sit through it.

I wish to dedicate this paper to the memory of my dear friend John Macnamara, who died in January 1996. Over the many years I had the privilege of knowing John, his loving concern with plumbing the depths of seemingly innocent questions of psychology — and his deep moral sense that it *matters* to dwell on these questions — were an inspiration to me and to numerous other colleagues and students. He also was just one of my favorite people in the whole world, and my family's too.

This research was supported in part by a John Simon Guggenheim Fellowship to the author, in part by Keck Foundation funding to the Brandeis University Center for Complex Systems, an in part by National Science Foundation grant IRI 92-13849 to Brandeis University.

References

- Allport, A. 1989. "Visual attention". In M.I. Posner (ed), *Foundations of Cognitive Science*, Cambridge, MA: The MIT Press, 631-682.
- Baars, B. 1988. *A Cognitive Theory of Consciousness*. Cambridge: Cambridge University Press.
- Bellugi, U., Wang, P.P., and Jernigan, T.L. 1993. "Williams syndrome: An unusual neuropsychological profile. In S. Broman and J. Grafman (eds), *Atypical Cognitive Deficits in Developmental Disorders: Implications for Brain Function*. Hillsdale, NJ: Lawrence Erlbaum.
- Bickerton, D. 1990. *Language and Species*. Chicago: The University of Chicago Press.
- Block, N. 1995. "On a confusion about the function of consciousness. *Behavioral and Brain Sciences* 18: 227-287.
- Bregman, A.S. 1990. *Auditory Scene Analysis*. Cambridge, MA: The MIT Press.
- Broadbent, D.E. 1958. *Perception and Communication*. London: Pergamon.
- Carroll, L. 1895. "What the tortoise said to Achilles. *Mind* 4: 278-80 [Reprinted in Hofstadter (1979), 43-45].
- Cheney, D.L. and Seyfarth, R.M. 1990. *How Monkeys See the World*. Chicago: The University of Chicago Press.
- Crick, F. 1994. *The Astonishing Hypothesis: The Scientific Search for the Soul*. New York: Charles Scribners Sons.
- Culham, J.C. and Cavanagh, P. 1994. "Motion capture of luminance stimuli by equiluminous color gratings and by attentive tracking. *Vision Research* 34: 2701-2706.
- Dascal, M. 1995. "The dispute on the primacy of thinking or speaking. In M. Dascal, D. Gerhardus, G. Meggle, and K. Lorenz (eds), *Philosophy of Language — An International Handbook of Contemporary Research*, Vol. II, Berlin: De Gruyter, 1024-1041.
- Dennett, D.C. 1991. *Consciousness Explained*. New York: Little, Brown.
- Dennett, D.C. 1995. *Darwin's Dangerous Idea*. New York: Simon & Schuster.
- Edelman, G.M. 1992. *Bright Air, Brilliant Fire*. New York: Basic Books.
- Fodor, J.A. 1975. *The Language of Thought*. Cambridge, MA: Harvard University Press.
- Fodor, J.A. 1983. *Modularity of Mind*. Cambridge, MA: The MIT Press.
- Herskovits, A. 1986. *Language and Spatial Cognition*. Cambridge: Cambridge University Press.
- Hofstadter, D.R. 1979. *Gödel, Escher, Bach: An Eternal Golden Braid*. New York: Basic Books.
- Jackendoff, R. 1987. *Consciousness and the Computational Mind*. Cambridge, MA: The MIT Press.
- Jackendoff, R. 1992. *Languages of the Mind*. Cambridge, MA: The MIT Press.
- Jackendoff, R. 1996. "Conceptual semantics and cognitive semantics". *Cognitive Linguistics* 7: 93-129.
- Julesz, B. 1971. *Foundations of Cyclopean Perception*. Chicago: The University of Chicago Press.
- Köhler, W. 1927. *The Mentality of Apes*. London: Routledge & Kegan Paul.

- Lackner, J. 1985. "Human sensory-motor adaptation to the terrestrial force environment". In D.J. Ingle, M. Jeannerod, and D.N. Lee (eds), *Brain Mechanisms and Spatial Vision*. Dordrecht: Martinus Nijhoff.
- Landau, B. and Jackendoff, R. 1993. "'What' and 'where' in spatial language and spatial cognition". *Behavioral and Brain Sciences* 16: 217-238.
- Lerdahl, F. and Jackendoff, R. 1983. *A Generative Theory of Tonal Music*. Cambridge, MA: The MIT Press.
- Levinson, S. 1996. "Frames of reference and Molyneaux's problem: Cross-linguistic evidence". In P. Bloom, M. Peterson, L. Nadel, and M. Garrett (eds), *Language and Space*, Cambridge, MA: The MIT Press.
- Marr, D. 1982. *Vision*. San Francisco: Freeman.
- Minsky, M. 1986. *The Society of Mind*. New York: Simon & Schuster.
- Pullum, G. 1991. *The Great Eskimo Vocabulary Hoax*. Chicago: The University of Chicago Press.
- Searle, J.R. 1992. *The Rediscovery of Mind*. Cambridge, MA: The MIT Press.
- Smith, N. and Tsimpli, I.-M. 1995. *The Mind of a Savant*. Oxford: Blackwell.
- Vandeloise, C. 1991. *Spatial Prepositions: A Case Study from French*. Chicago: The University of Chicago Press.
- Wittgenstein, L. 1953. *Philosophical Investigations*. Oxford: Blackwell.
- Wrangham, R., McGrew, W.C., de Waal, F.B.M., and Heltne, P.G. (eds). 1994. *Chimpanzee Cultures*. Cambridge, MA: Harvard University Press.